



This is an author-deposited version published in: <http://oatao.univ-toulouse.fr/>
Eprints ID: 4945

similar papers at core.ac.uk

provided by C

To cite this document: NGUYEN Anh-Dung, SENAC Patrick, RAMIRO Victor and DIAZ, Michel. Pervasive intelligent routing in content centric delay tolerant networks. In: *The 9th IEEE International Conference on Pervasive Intelligence and Computing (IEEE PICom)*, 12-14 Dec 2011, Sydney, Australia.

Any correspondence concerning this service should be sent to the repository administrator:
staff-oatao@inp-toulouse.fr

Pervasive Intelligent Routing in Content Centric Delay Tolerant Networks

Anh-Dung Nguyen^{*†}, Patrick Sénac^{*†}, Victor Ramiro[‡] and Michel Diaz[†]
anh-dung.nguyen@isae.fr, patrick.senac@isae.fr, vramiro@niclabs.cl, michel.diaz@laas.fr

^{*}ISAE/University of Toulouse, Toulouse, France

[†]LAAS/CNRS, Toulouse, France

[‡]NIC Chile Research Labs, Santiago, Chile

Abstract—This paper introduces a Swarm-Intelligence based Routing protocol (SIR) that aims to efficiently route information in content centric Delay Tolerant Networks (CCDTN) also dubbed pocket switched networks. First, this paper formalizes the notion of optimal path in CCDTN and introduces an original and efficient algorithm to process these paths in dynamic graphs. The properties and some invariant features of these optimal paths are analyzed and derived from several real traces. Then, this paper shows how optimal path in CCDTN can be found and used from a fully distributed swarm-intelligence based approach of which the global intelligent behavior (i.e. shortest path discovery and use) emerges from simple peer to peer interactions applied during opportunistic contacts. This leads to the definition of the SIR routing protocol of which the consistency, efficiency and performances are demonstrated from intensive representative simulations.

Index Terms—Delay Tolerant Networks, Content-Centric Networks, Routing Protocol, Swarm Intelligence.

I. INTRODUCTION

With billions of increasingly processing and networking efficient nodes at the edge of the Internet, the question to use the fabulous capacity of this "peripheral computing, storage and networking cloud", in complement and independently of the traditional communication infrastructure and services offered by the legacy Internet, makes sense. The efficient use of this pervasive huge cloud of processing, storage and communication resources, potentially available "for free" but underlaid by security, power consumption and performance issues, has lead to a thread of researches in new communication paradigms and protocols adapted to this type of spontaneous and highly dynamic ad-hoc networks. Specially, two new complementary communication paradigms play a key role in this specific type of communication context, that is, pocket switched [6] and content centric communication [8]. Indeed, according to the dynamic topology, intensive churn, potentially huge space of nodes and nodes diversity intrinsic to this type of networks, classical end to end communication paradigms cannot be applied anymore. This leads to consider these networks as a source of content and to replace classical address resolution and management mechanisms by the pivotal concept of content id. In such networking context, the end to end paradigm can be profitably replaced by the publish/subscribe paradigm. Moreover, considering that end to end connectivity between content providers and users cannot be insured, the store-carry

and forward paradigm intrinsic to delay tolerant networks must be applied in this type of networks. This paper aims to contribute to the definition and design of new communication architectures, protocols and mechanism adapted to content centric delay tolerant networks (CCDTN) by proposing a new routing paradigm and protocol for the efficient routing of information between content providers and content users in pocket switched networks. Indeed, this original protocol aims to keep track and to follow optimal routes in CCDTN. This protocol, called Swarm Intelligence based Routing (SIR), rests on a "swarm intelligence based approach" of which the collective routing "intelligent" behavior results from the conjunction of nodes' simple individual fully decentralized behaviors. In this paper we push further the preliminary simulation based results introduced in [11].

The rest of this paper is structured as follows. In Section II, we formalize the notion of optimal path in DTN, III analyze several representative real traces to exhibit properties and invariant of these optimal paths. In Section IV, we introduce Swarm-Intelligence based Routing (so called SIR) protocol. Via simulation results, Section V demonstrates the consistency, efficiency and performances of SIR. Section VI reviews the most salient contributions on the issue of routing in DTN. Finally, Section VII concludes the paper.

II. FINDING SHORTEST PATHS IN DTNS

A. Definition

In a static graph, the classical shortest path problem consists in finding a path between two nodes such that the sum of the weights of its constituent edges is minimal. This definition of shortest path is adapted to static networks but cannot be applied to DTNs and dynamic networks in which links between nodes vary in time. A more suitable model for this type of network are temporal graphs [14], i.e. graphs in which a link represents an opportunistic contact at a given instant between two nodes. A path in a temporal graph can be seen as an ordered set of temporal links that allow a message to be transferred using the store-move-forward paradigm between two nodes. Formally, let l_{ij}^t be a link between node i and node j at instant t . A path from A to B is described by a time ordered set $l_{Ai}^{t_0}, l_{ij}^{t_1}, \dots, l_{kB}^{t_d}$, with $t_{i+1} > t_i$. Note that this definition is asymmetric (i.e. the existence of a path from A to B doesn't mean there exists a path from B to A). This

is due to the temporal order (i.e. appearance) of the links that constitute a dynamic path. We consider the two following metrics on dynamic paths :

- Delay : the sum of the inter-contact times between consecutive links constituting the path we consider that the communication delay is neglectable with regard to inter-contact delays).
- Number of hops : the number of temporal links which constitutes the path. This metric is related to the ressources (memory, communication, processing, energy...) used by the dynamic path.

Let $p_{AB}^{d,h}$ be a path which allows a message from a node A to reach its destination B in d time units and via h hops. From this point of view, two definitions of optimal path can be derived, depending on the optimization objective. For delay constrained communication the optimal path is the path giving the minimum amount of delay. If there are several paths giving the same delay, then the one giving the minimum number of hops is selected. On the resource constrained communication the optimal path could be the path giving the minimum number of hops. In this case, if there are several paths with the same number of hops, the one giving the minimum delay is selected. In the following, we will focus on delay-constrained path only, in this case the delay is also called the path length.

B. Pairwise Shortest Paths Algorithm

The pairwise shortest path problem consist of finding the shortest path between every pairs of vertices in a graph. We propose here an algorithm to solve this problem in a temporal graph with the notion of shortest path as defined in the previous section. This algorithm leverages on the property of the adjacency matrix in static graphs.

In graph theory, the adjacency matrix is defined as the matrix A in which the element $a_{ij} \in \{0,1\}$ denotes the existence of a link between node i and node j . This matrix has the following interesting property : if we process the power n of matrix A , $B=A^n$, then its element b_{ij} gives the number of paths of length n between i and j . Indeed, for example, when $n = 2$, $b_{ij} = \sum_k a_{ik} \times a_{kj}$ sums all the possibilities to go from i to j through an intermediate node k .

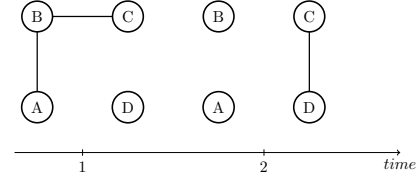
This property can be easily extended to temporal graph. Let $A(t), t = 0, 1, \dots, n$ be the adjacency matrix of a temporal graph at time t . The matrix $C(t)$ obtained as follows

$$C(t) = A(t)_b \wedge A(t)_b^2 \wedge A(t)_b^3 \dots \wedge A(t)_b^n, \quad (1)$$

where $A(t)_b^i$ denotes the binary version of the matrix $A(t)^i$ (i.e. the element $a(t)_b$ equals to 1 if $a(t) > 0$ and 0 otherwise) has its elements $c_{ij}(t) \in \{0,1\}$ that indicate if there is a direct or indirect (i.e. via multiple hops) link between i and j at instant t . Indeed, $c_{ij}(t)$ is the logical sum of all possibilities to have a direct or indirect link (up to n hops) between i and j at time t .

In consequence, the product

$$D(t) = C(0) \times C(1) \dots \times C(t) \quad (2)$$



(a) Illustration of a temporal graph

$$\begin{aligned} A(1) &= \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, A(2) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \\ \Rightarrow C(1) &= \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, C(2) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \\ \Rightarrow D(1) = C(1) &= \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \\ D(2) = C(1) \times C(2) &= \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \\ \Rightarrow L &= \begin{bmatrix} 0 & 1 & 1 & 2 \\ 1 & 0 & 1 & 2 \\ 1 & 1 & 0 & 2 \\ \infty & \infty & 2 & 0 \end{bmatrix} \end{aligned}$$

(b) Shortest path length computation

Fig. 1. Example of all-pairs shortest paths algorithm

results in a matrix in which the element $d_{ij}(t)$ specifies, when not null, that there is temporal path of delay t between i and j . As a consequence, the shortest path length from node i to node j is given by the smallest value of t such as $d_{ij}(t)$ equals to 1.

It is trivial to demonstrate that if a node C belongs to a shortest path between node A and node B then the AC sub path gives the shortest path between A and C . Therefore the shortest path between two nodes A and B can be easily backwardly reconstructed.

Finally, an example of the algorithm is illustrated in Figure 1. In this example, the final matrix L contains the shortest paths lengths of all pairs of nodes in the network.

In practice, as it's unlikely to have a large number of nodes connected to each other at a given moment, we can optimize the algorithm by limiting the number of iterations n in Equation 1 to an upper bound of the network diameter.

III. SHORTEST PATHS PROPERTIES IN DTNS

A. Real Traces Analysis

In this section, with the help of the algorithm previously introduced, we will study the evolution and properties of shortest paths that can be observed on real mobility traces. Indeed, the understanding of these properties is important to design efficient routing protocols adapted to DTNs. To study the properties of shortest paths in real DTNs, we analyzed the traces from the Haggle project [2] which consists of records

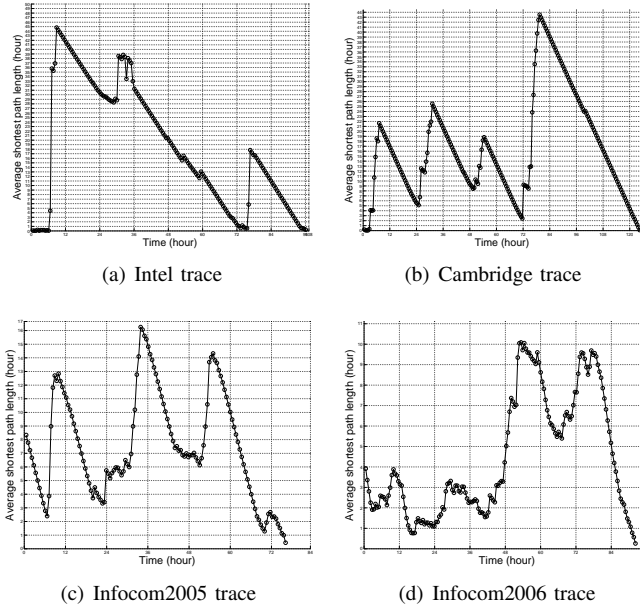


Fig. 2. Evolution over time of the average shortest path length of real DTNs

of Bluetooth range contacts between devices held by attendees during a conference or within a group of students/researchers in a laboratory. First, temporal snapshots of the network state are taken and the related adjacency matrix is processed with a 200s periodicity. Then, the shortest paths between all pairs of nodes are computed. Three salient properties of the shortest path length can be observed from these traces.

1) *Periodic Pattern of the Shortest Path Length:* Figure 2 shows a typical example of the temporal evolution of the shortest path average length. This figure shows the periodic evolution of the average shortest path length observed in different traces. The approximately 24 hours period with two opposite phases results from night/day variations in human activities and the resulting high change in shortest path lengths. Indeed, during the day, peoples' alternate movements and meetings leading to shorter path length than during static night periods. On the contrary, at night, they are more static and have less chance to meet each other, leading to longer path lengths. An adaptive communication stack can leverage on this periodic behavior to switch, for instance, from a DTN oriented communication scheme to a point to point one during phases with low mobility and meeting rates.

2) *Symmetry of the Shortest Path Length:* As defined in Section II-A, paths in DTNs are generally asymmetric due to the temporal ordering of the links constituting the path. It is interesting to check on real traces if this property entails differences between the shortest path length of a path and the one simultaneously observed for the reverse path. In order to check this feature, on several traces, for each couple of nodes (n_i, n_j) we measured the shortest path length between n_i and n_j and simultaneously in parallel for the reverse path between n_j and n_i . Figure 3(a) shows a typical evolution over time of these two shortest path lengths between two nodes and

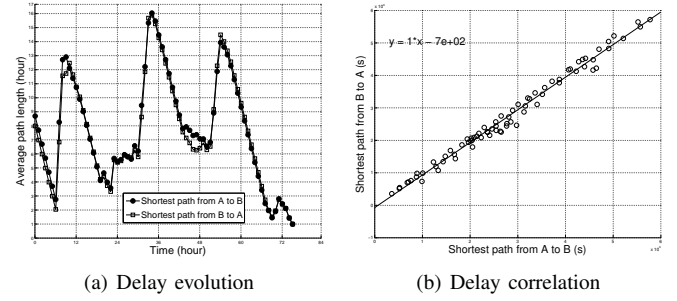


Fig. 3. Symmetry of the shortest paths in Infocom05 trace

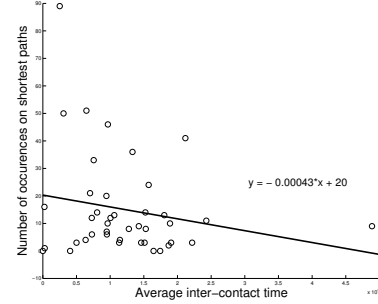


Fig. 4. The involvement of a node on the shortest path of another node is inversely proportional to the average inter-contact time between these two nodes

Figure 3(b) shows the correlation between the two time series of values with the related regression line. Interestingly, the result shows a nearly permanent symmetry of the shortest path lengths. In the CCDTN context, this symmetry entails that if the shortest path is built from content user to the content publisher and if we suppose some stationarity of this length (as observed in the traces during day phases) then the content can be delivered from the publisher to the user along a path with approximately the same length.

3) *Pairwise Inter-contact Time vs the Shortest Path:* The characterization of the nodes involved in shortest paths is an important issue. Indeed, an efficient routing algorithm should take advantage of this characterization to forward content to the best candidate nodes. Using our algorithm, we process, in the Infocom05 traces, the shortest paths between all nodes and a selected node A . Then, for each node, we count the number of its occurrences on the shortest paths to A . At the same time, we compute the pairwise inter-contact time between these nodes and A . Figure 4 shows the correlation between the average number of occurrences in the set of shortest paths and the average inter-contact time for all the nodes of the considered trace. This figure shows that the shorter is the inter-contact time between a node and the destination, the higher is its frequency of occurrence on the whole set of the shortest paths. This means, according to the spatiotemporal correlation observed in human mobility patterns [5], that shortest paths have a tendency to converge towards the nodes that the destination frequently meets.

IV. SWARM-BASED INTELLIGENT ROUTING (SIR)

Leveraging on the definition of shortest path in dynamic networks and the properties of this path introduced in the previous sections, we have introduced in [11] SIR, *Swarm-based Intelligent Routing (SIR)* a novel routing protocol inspired by the collective swarm intelligence that can emerge from basic individual behaviors applied by each element of a swarm of autonomous agents. In SIR, network nodes, via opportunistic contacts, maintain for each subscription they have heard of an utility function which sums up how close (as defined in Section II) this node is from a content subscriber. We will show that the distributed set of scalars associated to a given subscription forms across the network a gradient field in which the maximum value is carried by the content provider and the minimum value by the content user. Therefore content routing with SIR simply consists in following the steepest slope towards the minima of the gradient field where content users are located. Such algorithm works in a complete distributed way, so that nodes do not need to maintain any knowledge about the global topology of the network. This section, briefly introduces the SIR protocol of which a more detailed description can be found in [11].

Putting in synergy the DTN and publish/subscribe paradigms, SIR is composed of two phases.

A. Interest Dissemination Phase

In this phase, nodes register their interest for a content by disseminating efficiently this interest into the network. We leverage on this interest dissemination to dynamically establish a gradient field of which the intensity decreases from the content providers to the content users. The Binary-Spray-and-Wait protocol [15] well known to offer a good trade-off in term of delay, delivery ratio and resources use, is selected for the efficient diffusion of subscription messages. In SIR, a relaying node receives not only copies of subscription messages but also a couple of metric values defined in the previous section (i.e. delay and number of hops). SIR relay nodes update their utility values through opportunistic contacts according to their encounter's delays and number of hops from the content users. Hence, this behavior contributes to establish gradient fields that keep track of the shortest path between content producers and content users. Relay nodes utility is updated according to the following behavior:

- If a relay meet a content user, it resets its delay to 0 and updates its number of hops to 1.
- If two relays meet each other, the node with the higher delay sets its delay equal to the delay of the other node and its number of hops equal to the number of hops of the other node plus 1
- If two relays meet each other and have the same delay, then they keep their utility values unchanged.
- Otherwise delays progresses according to time evolution.

Figure 5 illustrates graphically this mechanism. On this figure, t and h are respectively the delay and the number of hops.

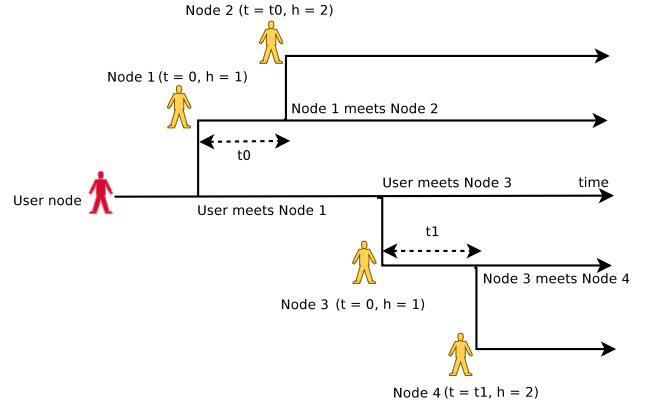


Fig. 5. Example of the interest dissemination phase

B. Content Dissemination Phase

When a content publisher receives the interest message, the second phase consists in sending back the content to the user by following, ideally, a path as close as possible to the shortest path between the content provider and the content user. The Binary Spray and Wait protocol is also used as the underlying diffusion mechanism. The content forwarding decision mechanism consists in always selecting a relay node of which the delay is smaller (i.e. a node closer in time to the content user). If the delays are equal, the content is forwarded if the encountered relay has a smaller number of hops (i.e. closer to the content user in space).

V. SIMULATION RESULTS

This section pushes further the preliminary results introduced in [11] by focusing on a more complete and in depth simulation-based evaluation of the SIR routing protocol. First, we verify the hypothesis that the utility function synthesizes coherently the proximity of a node to a content subscriber. Secondly, we run the protocol on real traces and compare the delay of the paths found by SIR with the shortest path found by the algorithm given in Section IV. This makes it possible to assess how close paths taken by SIR are to shortest paths. Finally, we compare SIR performances with two classical DTN-oriented routing protocols : PROPHET and Spray-and-Wait.

A. Consistency and Efficiency Issues

1) *Consistency of the Utility Function*: As introduced in the previous section, the utility function aims to sum up the proximity of a node to a related content subscriber. Therefore, a positive correlation between these two parameters should entail a linear relationship between the utility value and the shortest path length. To verify this hypothesis, we have implemented the SIR algorithm in MATLAB and we ran it on Infocom2005 traces which record Bluetooth connectivity of 41 handheld devices carried by conference attendees during 3 days. In the selected scenario, one content user sends his interest for this content to one content publisher and then receives the periodically refreshed content from the later. 41 interest

0.9001	0.8785	0.9184	0.5741	0.8475	0.8717	0.8358	0.8663
0.8830	0.9234	0.8439	0.7360	0.8065	0.9063	0.0403	0.9203
0.9184	0.7001	0.9296	0.7816	0.8886	0.9265	0.8932	0.8285
-0.1854	0.7513	0.9327	0.7761	0.9192	0.4026	-1.0000	0.7319
0.9080	0.9221	0.9071	0.9655	0.8687	0.9670	0.6925	0.9420
0.0854							

TABLE I
VECTOR OF CORRELATION COEFFICIENTS FOR ALL NODES

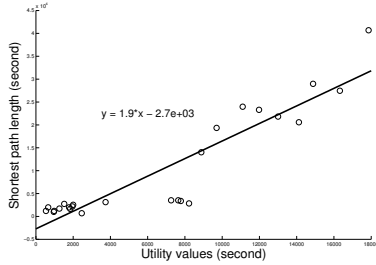


Fig. 6. Correlation between the utility values of a node and the shortest path length from that node to the content subscriber

message copies, initially disseminated by the subscribing node, allows one to establish a gradient field which is maintained and reinforced by all the nodes during the simulation. On one hand, we measure the evolution of nodes' utility values and on the other hand we process the shortest path length between the content subscriber and every other nodes. Then, we take the average value over 10 simulation runs. Figure 6 shows the correlation between these two time series. One can observe that the utility values and the shortest path length are highly correlated with 68% of correlation coefficient higher than 0.8, which reinforces our hypothesis.

2) *SIR vs Optimal Solution*: The goal of this evaluation is to show that by using the gradient field maintained and used by SIR one can achieve a nearly optimal solution. We conduct the same experiment described in the previous section. Then, we measure the content delivery delay at different instants of the day. This procedure is repeated with different pairs of user/publisher. On the other hand, we run our pairwise shortest paths algorithm to find the shortest path between these pairs of nodes. Figure 7 shows the average result over 10 pairs of nodes. These results show that SIR finds paths with delivery delays very close to the optimal paths.

B. Performance Issues

In this section, from a simulation-based point of view, we compare the SIR routing performances with the ones given by a basic Binary Spray and Wait (BSW) protocol and the ones delivered by PROPHET - a probabilistic based routing protocol. The performances delivered by the BSW routing protocol have to be considered as floor performances that SIR, by extending BSW with an utility function, should overtake in every situation. PROPHET can be considered as a swarm inspired protocol based on the distributed processing of an inter-node encounter probability. Therefore, the comparison of SIR and PROPHET performances allows one to classify our approach in comparison with a probabilistic approach

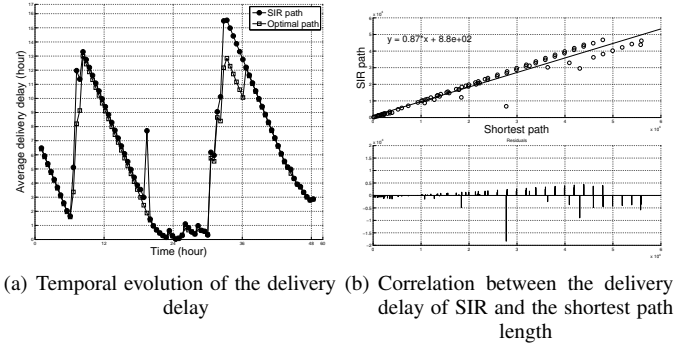


Fig. 7. Delivery delay of SIR vs optimal solution

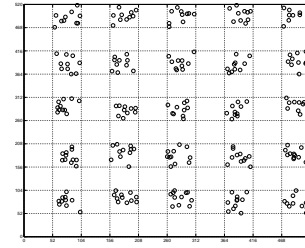


Fig. 8. Initial spatial distribution of nodes in STEPS simulations

known to perform well in DTN networks. Although Binary Spray and Wait and PROPHET were designed according to an end-to-end communication paradigm, these protocols can also be adapted to pub/sub systems with address based routing replacing content routing.

In order to reproduce at the simulation level realistic human mobility patterns, we have used the STEPS mobility model and simulator [10]. As we have shown in a previous work, this flexible parametric model can express a large spectrum of mobility patterns: from highly nomadic ones to localized ones; allowing us to evaluate routing protocols in different mobility contexts. Moreover, this model was shown to reflect accurately at the simulation level the routing performances observed with real traces. In the following scenarios, the network area is modeled as a torus divided in a number of square zones. Inside these zones, nodes move following the Random Waypoint model. Each node is attached to one preferential zone. The movement of nodes between these zones is driven by a parameter of the STEPS model which allows the nodes nomadism to be enforced or reduced according to a power law distribution (i.e. the probability that a mode moving outside his preferential zone has to return to that zone). Figure 8 illustrates the initial nodes' spatial distribution and Table II summarizes the main simulation parameters.

The communication protocol stack used in these simulations is simplified and focuses mainly on the network layer. For the sake of simplicity, we assume that the connections between nodes have an infinite bandwidth and that nodes have an infinite buffering capacity. The communication channel is modeled as a unicast channel. When two nodes are in communication range, they exchange instantly their buffers

TABLE II
SIMULATION CONFIGURATIONS

Number of nodes	100
Number of content subscriber	1
Number of content provider	1
Network size	$300 \times 300 \text{ m}^2$
Number of zones	10×10
Duration	10000 s
Warm up	2000 s
Radio range	10 m
Node speed	[3, 5] km/h
Content release frequency	100 s
Number of content copies (for SIR and BSW)	10

containing interests, contents and utilities.

We chose to simulate 3 different scenarios to study the impact of three fundamental factors on the protocols: mobility, connectivity and scalability. In these scenarios, we use the average delivery delay and the average delivery ratio (i.e. the ratio between the number of messages received and the number of messages sent) as performance measures.

1) *Scenario A - Impact of Mobility*: Mobility is known to have a significant impact on routing performances. In particular, in Pocket Switched Networks, human mobility plays an important role in routing protocol design because it contributes to packets mobility and packet routing. We exposed the three protocols to two different mobility contexts. The first one corresponds to a "sedentary" mobility behavior where people tend to spend the main part of their time in specific locations (i.e. home, workplace...) and move barely far from these preferential places. In STEPS, this characteristic is expressed with nodes having a strong attraction towards their preferential zones. The second mobility context corresponds to a more "nomadic one", such as human movements during a social event or in shopping areas. This behavior is modeled with nomadic nodes that are weakly attached to any preferential zone and tend to move quite randomly in the simulation area.

Simulating each protocol on these two mobility contexts, we measured the content delivery delay and the content delivery ratio between different pairs of nodes. In our simulations, the same mobility trace was used to ensure that the different protocols run exactly on the same mobility scenario. In each scenario, contents are sent periodically by publishers to the subscriber every 100 seconds. Then, we measured the average delivery delay and average delivery ratio between different pairs of nodes. The final result is averaged over 10 simulation runs. Figure 9 shows the bar plot of these measures. The segment on top of each bar represents the 95% confidence interval of the estimated mean value.

These results show that in "sedentary" mobility context, SIR outperforms the other protocols by significantly decreasing delivery delay and increasing delivery ratio. It is interesting to see that even in the "nomadic" context, a-priori less favorable to SIR which leverages on spatiotemporal correlations, SIR still slightly outperforms the other protocols. On the other hand, as intuitively understandable, for the three

TABLE III
SCENARIO C SETTINGS

Scenario	Number of Nodes	Network Size
C1	100	$300 \times 300 \text{ m}^2$
C2	200	$430 \times 430 \text{ m}^2$
C3	300	$520 \times 520 \text{ m}^2$

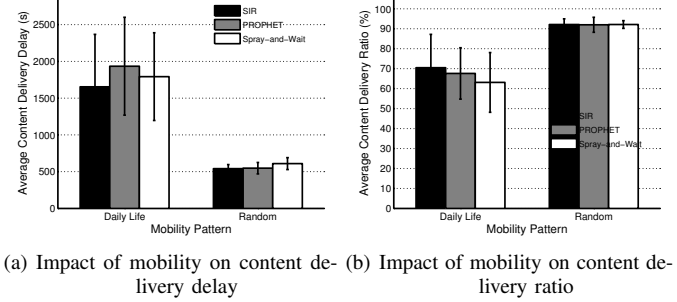


Fig. 9. Scenario A - Routing performances of the three routing protocols under different mobility contexts

protocols, the nomadic context contributes to increase routing performances and entails much higher delivery ratio and lower delivery delay than the more "sedentary" scenario. In the following scenarios, for comparison purpose, we will use the "sedentary" mobility context only, which conform to real life mobility patterns observed in real traces.

2) *Scenario B - Impact of Connectivity*: In this scenario, we perform the same experiment as described above but instead of varying the mobility pattern, we vary the connectivity between nodes by increasing their radio range from 10 m to 30 m. Again, the average delivery delay and average delivery ratio are measured for each scenario and for each protocol. Figure 10 gives the results obtained from these measures and show that SIR still outperforms the other protocols in all scenarios. Moreover, in the scenario with long range connectivity, which reduces significantly path lengths for all the protocols, SIR delivers still better performances both in term of delay and delivery ratio.

3) *Scenario C - Scalability*: Scalability is a very important feature specially for routing protocols that target potentially vast community of mobile users. In this scenario, we aim to test the capacity of SIR to route content efficiently in large-scale networks. For this, we increase the number of nodes in the network while keeping node density constant. The configurations for these scenarios are summarized in Table III. As seen on Figure 11, SIR scales well compared to the two other routing protocols, especially to Binary Spray and Wait. This can be intuitively explained by the fact that in large network, without any supplementary information and with a limited number of content copies, diffusion based protocols such as Binary Spray and Wait fail to relay content to good candidates which would allow the content to reach the destination with a good delay. On the contrary, SIR, with the additional routing constraint based on nodes spatiotemporal distance from the destination, improves significantly the routing performances.

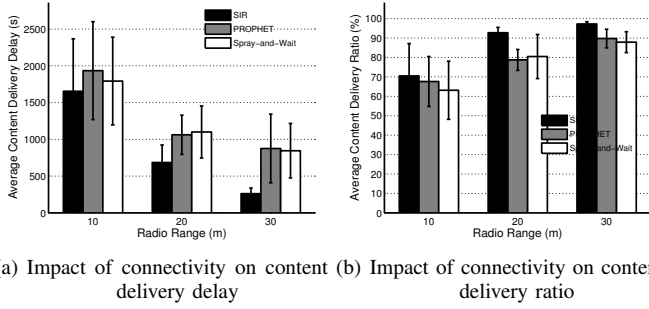


Fig. 10. Scenario B - Routing performances comparison of the three routing protocols under different connectivity levels

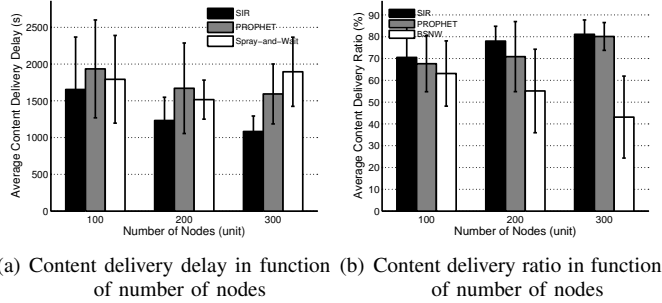


Fig. 11. Scenario C - Scalability of the three routing protocols

VI. RELATED WORKS

Basically, routing protocols for DTNs can be divided in two categories : non-context and context-based (see [3] for a survey). The first category consists in protocols in which nodes make forwarding decision while ignoring the network context (i.e. mobility or social structure). Protocols falling into this category belong to flooding and constrained-flooding families such as Epidemic Routing [16] or Spray-And-Wait [15]. In the second category, nodes exploit the context information (e.g. frequency of contact) from their local interactions to find the good candidate to forward the message. PROPHET [9], Bubble Rap [7], HiBop [1] and Propicman [12] fall into this category of routing protocols. In PROPHET, an utility value based on the contact frequency history is used to predict the probability of contact with the destination node. Bubble Rap makes the assumption of a community structure between nodes; messages are first pushed up to the most popular node in the community before being sent to the destination community and delivered to the destination node. HiBop and Propicman are more flexible for they do not make any assumption of a social structure. Nodes learn dynamically this structure via their local interactions. However, these protocols are deeply rooted in the end-to-end paradigm and cannot easily be adapted to the content-centric communication.

Attempts to bring the content-centric paradigm to opportunistic networks include SocialCast [4] and TACO-DTN [13]. In SocialCast, utility values based on social patterns and mobility are used for message forwarding decision. The evolution of these utility values is predicted with the help of a Kalman filter. TACO-DTN uses temporal information to

predict the future contacts between content providers and content users. This prediction is then used to schedule the delivery of messages. By resting on two universal and simple metric values intrinsic to the evolution of every networks (i.e. space with the number of hops and time with the delay), SIR distinguishes from these protocols based on more complex and artificial interaction patterns.

Finally, the notion of temporal graph was first addressed in [14]. In this work, the authors formalize the notion of temporal path length, temporal clustering coefficient and temporal efficiency to study properties of temporal graph. We leveraged of this definition of temporal path length to formalize the notion of shortest path in DTN.

VII. CONCLUSION

From a formal definition of the notion of shortest paths in dynamic networks and an in depth analysis of shortest path properties observed in representative real traces, we have introduced in this paper, SIR, a simple and light routing protocol that makes it possible to found and follow efficient routes in dynamic networks. The SIR protocol puts in synergy the DTN and CCN paradigms to found and follow shortest paths between content providers and content users in Delay Tolerant Content Centric Networks (DTCCN). We have shown in this paper that efficient and high performing routing can emerge from simple and light individual rules applied by each node during their opportunistic contacts. Simulations, from a parametric mobility model that makes it possible to cover a broad scope of mobility patterns, show that the proposed protocol, leads to the distributed elaboration of consistently decreasing gradient fields between content users and content providers. Therefore, routing with SIR simply consist in following steepest slope between content providers and content users. These simulations have also shown that SIR is never outperformed by classical information diffusion mechanisms such as Binary Spray and Wait and probabilistic routing. Moreover, SIR increasingly outperforms these protocols when the spatiotemporal correlation between nodes increases. According to the strong spatiotemporal correlation frequently observed in real mobility traces, SIR is a promising routing protocol for content centric DTN. However, we are convinced that there is still a margin of progress for SIR, specially by introducing, for example, inter-contact and intra-contact delays experienced by each node for the processing and actualization of the utility metric, and by studying the scalability and resource use in response to very large nodes and data spaces.

REFERENCES

- [1] Chiara Boldrini, Marco Conti, Jacopo Jacopini, and Andrea Passarella. Hibop: a history based routing protocol for opportunistic networks. In *IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks*, 2007, pages 1–12, June 2007.
- [2] Augustin Chaintreau, Pan Hui, Jon Crowcroft, Christophe Diot, Richard Gass, and James Scott. Impact of human mobility on opportunistic forwarding algorithms. *IEEE Transactions on Mobile Computing*, 6(6):606–620, June 2007.
- [3] M. Conti, J. Crowcroft, S. Giordano, P. Hui, H.A. Nguyen, and A. Passarella. Routing Issues in Opportunistic Networks. *Middleware for Network Eccentric and Mobile Applications*, 1:121, 2009.

- [4] Paolo Costa, Cecilia Mascolo, Mirco Musolesi, and G.P. Picco. Socially-aware routing for publish-subscribe in delay-tolerant mobile ad hoc networks. *IEEE Journal on Selected Areas in Communications*, 26(5):748–760, June 2008.
- [5] Marta C González, César A Hidalgo, and Albert-László Barabási. Understanding individual human mobility patterns. *Nature*, 453(7196):779–82, June 2008.
- [6] P Hui, A Chaintreau, J Scott, R Gass, J. Crowcroft, and C. Diot. Pocket switched networks and human mobility in conference environments. In *Proceedings of ACM/SIGCOMM WDTN 2005*, pages 244–251. ACM, 2005.
- [7] Pan Hui, Jon Crowcroft, and Eiko Yoneki. Bubble rap: social-based forwarding in delay tolerant networks. *IEEE Transactions on Mobile Computing*, pages 241–250, 2010.
- [8] Van Jacobson, DK Smetters, JD Thornton, M.F. Plass, N.H. Briggs, and R.L. Braynard. Networking named content. In *Proceedings of the 5th international conference on Emerging networking experiments and technologies*, pages 1–12, 2009.
- [9] Anders Lindgren, Avri Doria, and O. Schelén. Probabilistic routing in intermittently connected networks. *ACM SIGMOBILE Mobile Computing and Communications Review*, 7(3):19–20, 2003.
- [10] A.D. Nguyen, Patrick Sénac, and Victor Ramiro. STEPS-an approach for human mobility modeling. In *NETWORKING 2011*, pages 254–265. Springer, 2011.
- [11] A.D. Nguyen, Patrick Sénac, and Victor Ramiro. Swarm-based Intelligent Routing (SIR) - a New Approach for Efficient Routing in Content Centric Delay Tolerant Networks. In *ACM MOBIWAC 2011*. ACM, 2011.
- [12] H.A. Nguyen, Silvia Giordano, and Alessandro Puiatti. Probabilistic routing protocol for intermittently connected mobile ad hoc network (propicman). In *IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks, 2007*, pages 1–6, June 2007.
- [13] Giuseppe Sollazzo, Mirco Musolesi, and Cecilia Mascolo. TACO-DTN: a time-aware content-based dissemination system for delay tolerant networks. In *Proceedings of the 1st international MobiSys workshop on Mobile opportunistic networking*, pages 83–90, June 2007.
- [14] T Spyropoulos, K Psounis, and C.S. Raghavendra. Efficient routing in intermittently connected mobile networks: the multiple-copy case. *IEEE/ACM Transactions on Networking (TON)*, 16(1):77–90, 2008.
- [15] Thrasyvoulos Spyropoulos, Konstantinos Psounis, and C.S. Raghavendra. Spray and wait: an efficient routing scheme for intermittently connected mobile networks. In *Proceedings of the 2005 ACM SIGCOMM workshop on Delay-tolerant networking*, pages 252–259, 2005.
- [16] A Vahdat, D Becker, and Others. Epidemic routing for partially connected ad hoc networks. Technical Report CS-200006, 2000.